Disaster Recovery in Cloud Computing

Mr.Akshay A. Gharat, Mr. Devendra E. Mhamunkar

ASM INSTITUTE OF MANAGEMENT & COMPUTER STUDIES (IMCOST), THANE, MUMBAI

University Of Mumbai

Abstract: Nowadays , data has been generated in large amount that required the data recovery services. We know that the cloud computing introduces a new type of computing platform in today's world. This type of computing will generates a large amount of private data on main cloud. Therefore, the necessity of data recovery services are growing day-by-day and it requires a development of an efficient and effective data recovery technique. The purpose of recovery technique is to help user to collect information from any backup server when server lost his data and unable to provide data to the user. To achieve this purpose, many different techniques have been proposed till date. In this review paper, we mention few recent techniques that are the solutions in the form of "Disaster Recovery Techniques and Online Data Backup ". The objective of this review paper is to summarize the powerful data backup recovery techniques that are used in cloud computing domain.

Keywords: **Cloud Computing, Disaster recovery techniques, Traditional disaster recovery, Disaster recovery planning, Disaster recovery as a service, Backup, Privacy.**

I. INTRODUCTION

To handle devices or licensed software's in an organization having more employees and to provide those software's to all employees for their work without any delay is slight difficult by using physical hardware. To overcome this problems cloud computing is developed. Cloud computing is internet based computing process in which systems are interconnected with sharing resources by each other. Internet is the medium acts between cloud and user. Client is connected to cloud server and can store data through internet and can access the data from anywhere. it is a real time communication network. we can run our programs from anywhere by accessing cloud. using cloud we can access any software or data without paying any money to cloud.

When a system crashes or power failure occurs there is a chance of loss of data and sometimes it may result it in financial loss. This system crashing and other problems occur due to natural Disasters or by human from causing expensive service disruptions. When a disaster occurs in business continuity the company may get huge loss of data and also financial loss. When disaster occurs company need to protect the data from loss. Cloud providing companies like Google, Amazon, Microsoft etc., experienced cloud disaster with a huge loss of data and servers. When disaster occurs at client side backup will be stored in cloud but if disaster occurs in cloud data will be lost. Natural disasters may occur due to bad weather results in disaster To overcome these disasters there are some disaster recovery techniques which are used to recover data.

Disaster recovery techniques as required to their business continuity. Dedicated and shared models are the two approaches for disaster recovery based on cost and speed. Storing the data from cloud infrastructure in order to recover when disaster occur. Every organization should have a documented disaster recovery process and should test that process at least twice each year.

II. CAUSES OF DATA LOSS

A. Natural Disasters

When natural disasters occur then 2% of the data will be lost. The main reasons of occurrence natural disasters are mundane and nefarious. Due to mundane and nefarious effects one cannot recognize the data loss when disaster occurs.

B. Mission critical application failure

When an application is left unusable for few days then it causes a sudden great damage failure and in some organizations it may be mission critical. By using all applications that are stored in cloud may reduce the sudden great damage.

C. Network failure

Cloud and clients are connected by internet and when network fails the systems which are connected to cloud are crashed and data will be lost and applications which are working based on cloud will also suffer.

D. Network intrusion

When a virus is invaded onto the applications then there is a chance of occurrence of disaster. By placing unusable applications in that place on a watch list we can prevent occurrence of disaster.

ISSN: 2278 – 1323

E. Hacking or malicious code

Disaster occurs in inside or outside of the organization although they prevent hacking or malicious code from modifying data there is a loss of data.

F. System failure

If infrastructure in a organization fails then whole systems which are connected in that organization will crash. This will affect the operating systems.

The main reason for occurrence of disaster is human, 60% of the data centres are failed.

III. TRADITIONAL DISASTER RECOVERY

Traditional disaster recovery was developed by share group which are divided into 6 tiers.

A. Tier 0

No offsite data that means there is no disaster recovery plan and no saved data. To recover data it may take weeks and it is unsuccessful.

B. Tier 1

Data backup without hotsite that means data is taken backup by offsite not by hotsite. To retrieve the data that is taken backup is time taken process. By not having their own redundant servers it is time taking process to locate and configure appropriate systems.

C. Tier 2

Data backup with hotsite that means organizations maintain data backup as well as hotsite it is the fastest process. By having a hot backup site when disaster occurs we can run applications at stand by servers.

D. Tier 3

Instead of taking backup by physical media it provides an electronic vault so that backup data is network accessible to hotsite. As hotsite backup is cost effective it is better to access it by network.

E. Tier 4

Point in time copies means that organization maintains more timely point in time backup of crucial data is network accessible to host site.

F. Tier 5

Transaction integrity means that transactions are consistent between production systems and recovery sites. So, there should be no loss of data.

IV. DISASTER RECOVERY REQUIREMENTS

This explains key features for effective cloud service when disaster occurs.

A. Recovery point objective

Maximum time period taken for data loss when an disaster occurs is calculated RPO. The necessary RPO is generally a business decision—for some applications absolutely no data can be lost (RPO=0), requiring continuous synchronous replication to be used, while for other applications, the acceptable data loss could range from a few seconds to hours or even days. The recovery point objective identifies how much data you are willing to lose in the event of a disaster.

Your RPO is typically governed by the way in which you save and back up data:

• Weekly off-site backups will survive the loss of your data centre with a week of data loss. Daily off-site backups are even better.

• Daily on-site backups will survive the loss of your production environment with a day of data loss plus replicating transactions during the recovery period after the loss of the system. Hourly on-site backups are even better.

• A clustered database across multiple data centres will survive the loss of any individual data centre with no data loss.

B. Recovery time objective

It is a measurement of time upto which it can withstand and bring back to the system when a disaster occurs. It may be minutes, hours, and days. It may also include detection of failure and preparing required servers at backup site to initialize an application which is interrupted in middle of execution. The recovery time objective identifies how much downtime is acceptable in the event of a disaster.

V. DISASTER RECOVERY PLANNING

There are some mechanisms that are implemented for data backup when disaster recovery technique is used. So that when we want to take backup of a data we can follow some mechanisms.

Backup sites can come from three different sources:

- Companies specializing in providing disaster recovery services.
- Other locations owned and operated by your organization.
- A mutual agreement with another organization to share data centre facilities in the event of a disaster.

A. Hot Backup Site

It is very expensive to operate. This site works with organizations that operate real time processes.

ISSN: 2278 – 1323

It is the duplicate of the original site. Loss in data is very minimal as we can relocate the data and continue our work what we are performing. It will save as a virtual image of our current data. In a few hours hot backup site can bring up to full production. It is priory used in the situations where disaster happening.

B. Cool backup site

It is the least expensive to operate. It doesn't take any backup of data copies or it doesn't include hardware. Lack of hardware can start-up with a minimal cost but require more time. Everything required to restore service to users must be procured and delivered to the site before recovery operation is performed.

C. Warm backup site

It is already stocked with a hardware configuration on the backup site that found in primary site. To apply warm backup site the last data backup should be delivered to their primary sites.

VI. DISASTER RECOVERY AS A SERVICE

Disaster recovery as a service is an upcoming service as a nomenclature of cloud computing. It is a low cost service when compared to traditional disaster recovery. It is flexible in replicating physically or virtually. It provides application consistent recovery for some working applications like SQL server. It has pre-built options for virtual recovery environments including security, network connectivity and server failover when continuously replication among servers. When disaster occurs we can take backup and we can run our applications on service provided by disaster recovery until we get backup to primary site. Disaster recovery as a service to replicate critical servers and data centre infrastructure in cloud.

Fig. 1.Disaster recovery as a service

Disaster recoveries as a service is free or pay on use offer. When incompatibilities are occurred due to software changes then breaking of DRaaS in cloud may occur.

The architecture of DRaaS is defined by three models.

A. From Cloud

When the primary application or data is in cloud and backup or recovery site is in private data centre.

B. In cloud

When both primary site and recovery site are in cloud.

C. To cloud

When the application is in primary data centre and backup or recovery site is in cloud. To test the recovery processes sandboxes are used and they test without disrupting running application. It is only accessible to only system administrator.

Solutions are pre-packaged services that provide a standard DR Failover to a cloud environment that you can buy on a pay-per-use basis with varying rates based upon your recovery point objective (RPO) and recovery time objective(RTO).

VII. DISASTER RECOVERY CHALLENGES

In this section we investigate some common challenges of DR in cloud environments.

A. Dependency

One of the disadvantages of cloud services is that customers do not have control of the system and their data. Data backup is on premises of service providers as well. This issue makes dependency on CSPs for customers (such as organizations) and also loss of data because of disaster will be a concern for customers.

B. *Cost*

It is obvious that one of the main factors to choose cloud as a DR service is its lower price. So, cloud service providers always seek cheaper ways to provide recovery mechanisms by minimizing different types of cost. The yearly cost of DR systems can be divided in three categories (Alhazmi and Malaiya, 2012):

- Initializing cost: amortized annual cost
- Ongoing cost: storage cost, data transfer cost and processing cost
- Cost of potential disaster: Cost of recovered disasters and also cost of unrecoverable disasters.

ISSN: 2278 – 1323

C. Failure Detection

Failure detection time strongly affects on the system downtime, so it is critical to detect and report a failure as soon as possible for a fast and correct DR.

D. Security

As mentioned before, DR can be created by nature or can be human-made. Cyber-terrorism attack is one of human-made disasters which can be accomplished for many reasons. In this case, protection and recovery of important data will be a main goal in DR plans beside of system restoration.

E. Data Storage

By increasing of cloud usage in business and market, enterprises need to storage huge amount of data on cloud-based storages. In order to satisfy applications and also to n guarantee the security of data, computing has to be distributed but storage has to be centralized. Therefore, storage single point of failure and data loss are critical challenges to store data in cloud service providers

VIII. DISASTER RECOVERY SOLUTIONS

In this section, we will discuss some DR solutions which have been proposed to overcome the problems and challenges in cloud-based DR.

A. Local Backup

A solution for dependency problem has been proposed in (Javaraiah, 2011). A Linux box can be deployed on the side of customers to make control of data and to get backup of both data or even complete application. Local storage can be updated through a secured channel. By this technique, migration between cloud service providers and also migration between public to private, and private to public is possible. In the event of a disaster, local backup can provide the services that were served by the service provider.

B. Geographical Redundancy and Backup (GRB)

Although geographical redundancy can be used in traditional model, but it is expensive and unaffordable. In (Pokharel et al., 2010), two cloud zones have a replication of each other. If one zone becomes down, then another zone will be on and provide the services. There is a module that monitors the zones to detect disaster.

C. *Inter-Private Cloud Storage (IPCS)*

This approach was proposed for cloud data storage (Jian-hua and Nan, 2011). According to Storage Networking Industry Association (SNIA), at least three backup locations are necessary for business data storage. Users' data should be stored in three different geographical locations: Servers, Local backup server (LBS) and remote backup server (RBS). This model gives communication ability to backup locations in order to increase data integration.

D. Resource Management

Heterogeneous clouds consist many different hardware and software such as hybrid storage and diverse disks. In cloud-based enterprises, entire business data are stored in the cloud storage. So, data protection, safety and recovery are critical in these environments. Data in danger is the data which has been processed at the primary host but has not taken place in the backup host yet. So, in the case of disaster, It is necessary to use enhanced technology for data recovery in storage clouds. There are three solutions for data recovery proposed in (Patil et al., 2012)

- Using fastest disk technology in the event of a disaster for replication of data in danger.
- Changing dirty page threshold: The percentage of dirty pages in RAM which have to be waited for flushing to disk might be reduced (Rudolph, 1990).
- Prediction and replacement of risky devices: Some important factors such as power consumption, heat dissipation, carbon credit utilization and importance of data (stored on each disk) can be calculated in a specific period of time.

E. *Scale Up/Down*

Sometimes, performing functions with high priority can decrease money loss or even increase the revenue in the event of a disaster. After a natural disaster occurs in an area, cloud service providers are faced with flooding service requests. In this case, service providers have to manage their existent users' services and also handle new user requests. Service providers must satisfy existent users and should serve to new customers as much as possible. In (Nakajima et al., 2013), a management engine has been introduced for carrier networks. In case of a large scale natural disaster (like earthquakes), this system uses a DR scenario by scaling up resources for the high-priority services (e.g., voice communication) and scaling down allocated resources to low-priority service (e.g., video on-demand).

F. Use Of Storage Area Network To Eliminate Data Loss

The Issue: SQL Server database technology Enterprise edition offers an option - that of replicating changes to the primary database on a transaction by transaction basis to the database copy. The replication of a transaction, or small number of transactions, may not occur immediately after an interruption due to a disaster.

The small number of transactions occurring just prior to a disaster may not be captured in the replicated database is an issue for some companies and applications. The businesses

ISSN: 2278 – 1323

with stringent RPOs that are not satisfied with an asynchronous database replication solution can have their needs met with a synchronized approach. This requires a more effective technical solution described below.

Terminology
Replication can be done synchronously or asynchronously. Synchronous replication requires the application to wait until it knows that both copies, primary and backup, have been successfully written before proceeding. Asynchronous replication allows the application to proceed without confirmation under the assumption that the infrastructure will complete making the replicated copy. There is an exposure that this asynchronous replication activity can be interrupted at time of disaster, affecting the last few transactions.

The Technology: Storage Area Network (SAN) technology is a proven approach to ensure that data exists in identical form over a wide area network, generally using fiber channel protocol over an IP WAN (Internet Protocol, Wide Area Network). Its only drawback is cost. Cloud computing helps address the cost issue through scale economies. A number of quality vendors offer this functionality.

The Solution: The implementation involves SAN technology installed at both the in-house and virtual recovery locations and manages the replication of file and database data using the most advanced techniques. It should be noted that compatible SAN technology is required at both ends.
The SAN has the intelligence to manage the replication process and to communicate with applications regarding the completion of the update to the local and replicated copy. It is this intelligence that allows for a synchronous approach to replication, eliminating the risk of lost transactions.

Alternative Solution: The primary data centre is relocated from in-house into the cloud, operating on hosting vendor infrastructure that already has the SAN technology at its primary and secondary locations. The cost of using this approach would most likely be less than implementing a SAN in-house.

G. Pre-loaded Applications To Shorten Recovery Times

The Issue: Applications need to be installed, and software product keys entered. This can take time and requires involvement of someone with knowledge of both the software and the keys. Note that software vendors are not standardized on terms for allowing backup copies of software to be used: some allow a backup copy to be installed subject to rules of use; others do not.

The Technology: The virtual servers at the cloud hosting vendor site are operationally ready when needed and can be started from a remote site as needed.

The Solution: It is possible to pre-load application files at the cloud hosting site and enable them with keys at the time of disaster. If this solution is used, the change control process needs to include steps to ensure that any changes or updates to application software get replicated to the remote site immediately.

Solution 1: Pre-load applications and product keys, essentially making the applications ready to run if transactions are presented to them.

Solution 2: Some hosting vendors offer "tenant" software-as-a-service applications, fully supported and at lower cost than licensed versions. For example, MS Exchange can be licensed in the cloud, such that it is paid for only when in use.

IX. CONCLUSION

As cloud computing is becoming very important in day to day life and every company is based on cloud computing. They are not aware of disasters in cloud; they don't know any recovery mechanisms at first. When disaster occurred then all companies faced big loss of data and also financial then after many recovery mechanisms are introduced. As cloud nomenclature has a PaaS, IaaS, and SaaS as services which provide their service to cloud users in terms of infrastructure, software and platform as their requirement; so user can use cloud without any difficulty. By implementing DRaaS in cloud one can get recovered from data loss when he experiences a system failure or by natural disasters. So by implementing DRaaS in business continuity they can overcome their data loss.

REFERENCES

[1] https://en.wikipedia.org/wiki/Cloud_computing
[2] http://www.webopedia.com/TERM/C/cloud_computing.html
[3] http://www.cs.uwp.edu/staff/lincke/infosec/notes/BC-DR.ppt
[4] http://blog.ussignalcom.com/blog-1/bid/257525/6-Causes-of-Data-Loss-Prepare-your-Disaster-Recovery#_ftn2
[5] http://books.google.co.in/books?id=q0FaSaNEYK0C&pg=PA179&dq=cloud+disaster+recovery&hl=en&sa=X&ei=wPXUYCPJIWIrAfSzICABA&ved=0CEIQ6AEwAg#v=onepage&q=cloud%20disaster%20recovery&f=false
[6] http://en.wikipedia.org/wiki/Backup_site
[7] Cloud Application Architectures building applications and infrastructure in the cloud by O'RELLY
and George Reese.
[8] http://en.wikipedia.org/wiki/Recovery_as_a_Service
[9] https://www.usenix.org/legacy/event/hotcloud10/tech/full

ISSN: 2278 – 1323

_papers/Wood.pdf

[10] Aceto, G., Botta, A., Donato, W., & Pescape, A. (2013). Cloud monitoring: A survey. Computer Networks, 57(9), 2093-2915. http://dx.doi.org/10.1016/j.comnet.2013.04.001

[11] Javaraiah, V. (2011). Backup for cloud and disaster recovery for consumers and SMBs. *IEEE 5th International Conference on Advanced Networks and Telecommunication Systems (ANTS) (pp. 1-3).* http://dx.doi.org/10.1109/ANTS.2011.6163671

[12] Patil, S. R., Shiraguppi, R. M., Jain, B. P., & Eda, S. (2012). Methodology for Usage of Emerging Disk to Ameliorate Hybrid Storage Clouds. *IEEE International Conference on Cloud Computing in Emerging Markets (CCEM) (pp.1-5).* http://dx.doi.org/10.1109/CCEM.2012.6354615

[13] Pokharel, M., Lee, S., & Park, J. S. (2010). Disaster Recovery for System Architecture using Cloud Computing. *10th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT) (pp. 304-307).*

BIOGRAPHY

Akshay Ashok Gharat born in Alibag , Raigad district, Maharashtra India. He is pursuing MCA final year in ASM's Institute of Management & Computer Studies ,IMCOST, Thane

Devendra Eknath Mhamunkar born in Airoli, Thane district, Maharashtra India. He is pursuing MCA final year in ASM's Institute of Management & Computer Studies, IMCOST, Thane

ISSN: 2278 – 1323